DOCUMENT RESUME

ED 151 826 CS 204 009

AUTHOR	Wilks, Yorick
TITLE	Four Generations of Machine Translation Research and Prospects for the Future.
PUB DATE	77
NOTE	29p.; Paper presented at the NATO Symposium (Venice, Italy, September 25-October 1, 1977); Best available copy
EDRS PRICE	MF-$0.83 HC-$2.06 Plus Postage.
DESCRIPTORS	*Artificial Intelligence; *Computational Linguistics; Literature Reviews; *Machine Translation; Metaphors; Research; Research Problems; *Semantics

ABSTRACT

This paper begins with a description of four generations of research in machine translation: the original efforts of 1957 to 1965 and three types of surviving and sometimes competing present projects. The three types of present projects include those relying on "brute force" methods involving larger and faster computers; those based on a linguistic tradition which asserts that knowledge required for machine translation can be assimilated to the structure of a grammar-based system with a semantic component; and those stemming from artificial intelligence research, with an emphasis on knowledge structures. The paper argues that the artificial intelligence approach has the best chance of simulating the communicative abilities necessary for realistic machine translation and gives an account of how knowledge structures might cope with one of the classic problems of machine translation: that of metaphor, or "semantic boundary breaking." (AA)

Four generations of machine translation research and prospects for the future.

Yorick Wilks

Department of Language and Linguistics,

University of Essex,

Colchester, Essex, U.K.

Abstract

Although the large Government supported Machine Translation (MT) projects came to an end in the middle Sixties, it was by no means the case that all MT research stopped at that time. Some large projects continued in Universities (particularly in Canada), and some small commercial companies continued with the methods of the first generation, but using larger and faster computers. The results of their persistence have appeared only in the last few years. Moreover, research workers in Artificial Intelliegnce (AI) turned their attention to understanding natural language and, in some cases, to concrete problems of MT that required the manipulation of meaning and of real world information for their solution.

These latter three approaches are all alive and well at the moment, and the paper discusses their abilities to cope with the difficult problems of MT that were left unsolved by the first generation approaches. The bulk of the paper argues that for the long term future, the AI approach has the best chance of simulating the communicative abilities that will be required for realistic and general MT, and proceeds to give some account of how large scale knowledge structures might begin to cope with one of the classic problems of MT: that of metaphor, or "semantic boundary breaking".

Introduction

There is an ancient Chinese curse that dooms recipients to live in an interesting age, and by those standards Machine Translation (MT) workers are at present having a bad time. The reason things are interesting at the moment is that there is a number of conflicting claims in the air about how to do MT, and whether it can, or indeed has already, been done. Such a situation is unstable, and we may confidently expect some kind of outcome---always cheering for the empiricist---in the near future.

What has happened is threefold. First, the "brute force" methods for MT, that were thought to have been brought to an end by the ALPAC (1966) Report have surfaced again, like some Coelacanth from the deep, long believed extinct. Such systems are now being sold under such trade names as LOGOS, XYZYX, SMART and SYSTRAN; and the last, and best known, is now undergoing extensive testing in Paris (Etat 1976) and Luxembourg.

Secondly, some large-scale, more theoretically based, MT projects continued---usually based in Universities---and are now being tested in use, though sometimes on a scale smaller than that originally envisaged. METEO, for example, in Montreal (Chandioux, 1976), which was to have translated official documuments from English to French, is now in use for the translation of the more limited world of TV Weather reports.

Thirdly, workers in natural language in the field known as Artificial Intelligence (AI) have begun to make distinct claims about the need for their approach if there is to be general and high quality MT (Wilks, 1973; Charniak,

1973; Schank, 1975). Small pilot systems illustrating their claims have been programmed, but their role in contemporary discussion is mainly of a theoretical nature.

However, these are not merely three complementary approaches, for they seem to be making different claims, and, unless we take the easy way out and simply define some level of MT appropriate to each of the enterprises, it seems they cannot all be right, and that we may hope for some resolution before too long.

What I shall do in this brief paper is to sketch the recent background from the AI point of view, and then outline very briefly a development within the overall AI approach that should have some bearing on the possibility of high quality MT.

Some background notes

As the title hints and the introduction sets out, we now have, in my view, four generations of MT research: the original efforts of 1957-65 plus the three types of project now surviving, and indeed competing. The key to their relation can be found in their different responses to Bar-Hillel's critique of MT, which he updated at intervals, but which came down to one essential point: MT is not only practically but theoretically, impossible.

'Expert human translators use their background knowledge, mostly subconsciously, in order to resolve sytactical and semantical ambiguities which machines will either have to leave unresolved, or resolve by some

"mechanical" rule which will even so often result in a wrong translation. The perhaps simplest illustration of a syntactical ambiguity which is unresolvable by a machine except by arbitrary or ad hoc rules is provided by a sentence, say "... slow neutrons and protons....", have no difficulty in resolving the ambiguity through utilization of his background knowledge, no counterpart of which could possible stand at the disposal of computers." (Bar Hillel, 1962).

The immediate context of Bar Hillel's argument was the performance of early syntax analysers which, according to legend, were capable of producing upwards to of ten parsings of sentences like "Time flies like an arrow", where, with respect to standard dictionary information, any of the first three words could be taken as a possible verb.

The standard reaction to such syntactic results was to argue that this simply showed the need for linguistic semantics, so as to reduce the "readings" in such cases to the appropriate one. Bar Hillel's addition to this was to argue that it was not a matter of semantic additions at all but of the, for him unformalizable, world of human knowledge.

It is interesting to notice that the reactions of Bar Hillel and AI workers like Minsky were in part the same: (Minsky, 1968) argued that MT clearly required the formalization of human knowledge for a system that could be said to <u>understand</u>, or as Bar Hillel reviewed the situation in 1971 (Lehmann & Stachowitz, 1971, p.73),

"It is now almost generally agreed upon that high-quality MT is possible only when the text to be translated has been understood, in an appropriate sense, by the translating mechanism".

What Minsky and Bar-Hillel disagreed about, of course, was what followed: Bar-Hillel thought that the impossibility of MT followed, whereas Minsky believed that the task had now been defined, and the job of AI was to get on with it.

The contrast is clear between these two and the views of linguists: Chomsky's generative theories are also, in a clear sense, a reaction to the failure of early MT, in that they state with great force the case for a solid theory of natural languages as a precondition for any advance with machines and language. Fodor & Katz's semantics, adjoined to a generative grammar, represent, as it were, the linguistic analogue to those who thought that semantic information would resolve the mutliple parsings of the notorious "Time flies like an arrow". Later linguists broke from the Chomskyan paradigm by arguing that Fodor & Katz's rigid exclusion of human knowledge froma linguistic system was inadequate, and that many forms of pragmatic knowledge would be required in a full linguistic system. Lehmann - Stachowitz (1971) contains contributions along these lines from Ross and Fillmore, specifically in relation to MT.

The attempt by AI research to respond to Bar-Hillel's challenge is of a different sort. It is an attempt not only to admit ab initio the need for "knowledge-structures" in an understanding system, but also to formulate theories and systems containing processes for the manipulation of that knowledge. "Processes" here is not to be taken to mean merely programming a computer to carry out a task, for many AI systems of interest have either not been programmed at all or made to do only partial demonstrations. The

word "process" means that a theory of understanding should be stated in as symbol processing manner, one in which most linguistic theories are not information processing. This is a contentious position, in that generative grammar has been in some sense a description of a process since the earliest descriptions of transformational theory. The AI case is that it never quite comes up to scratch in processing terms. The nature of this dispute can be seen from such work as (Bresnan 1976) where an attempt is made to present transformational grammar at the highest level in an unfamiliar (to linguists) and process-orientated manner.

The METEO system represents what one might call the linguistic tradition in MT works: with claim that an MT system based on a linguistic theory is sufficient, and that whatever knowledge is required for MT can be assimilated to the structure of a grammar-based system with a semantic component.

The work of Ross and Fillmore referred to (as well by Lakoff and McCawley among others) represents a breakdown of the paradigm that has dominated linguistics since 1957, and in their search for more general notions of process to express knowledge computations it is no longer clear that anything fundamental separates them from what we have here called the AI approach to MT.

The additional contrast with the resurrected "brute force" methods should now be clearer. These approaches have in essence ignored the challenge of Bar-Hillel as well as the earlier one from linguistics for a theoretically

motivated syntax and semantics. The assumption behind work like SYSTRAN is that the main fault of the early MT period was inadequate machines and software, not theory. The striking demonstrations* given of that system are not yet conclusive, and detailed descriptions of its methods are not available because of understandable commercial considerations, but there can be no doubt that it does pose a considerable challenge to both linguists and AI theoreists, who claim, in their different ways that some higher level theory is essential for MT.

What then is an AI theory?

Apart from their common emphasis on knowledge structures and process form, AI theories can only be illustrated by example, since they differ so much among themselves on a wide range of issues (see Wilks, 1976c). Moreover, MT is not usually the implementation environment of a typical AI program (though see Wilks, 1973; 1975) which is normally dialogue, question-answering or paraphrase. But no issue of principle arises here, especially if one accepts Steiner's (1975) claim that every act of understanding is, in essence, one of translation.

Winograd's well-known program (1972) was perhaps the first AI language understander not directed to what one could call the classic residual problems of MT: word sense ambiguity, pronoun reprence ambiguity etc. He was concerned to show the role of knowledge of a microworld of blocks as a tool for resolving syntactic ambiguities in input to a dialogue system.

---

* One at the University of Zurich, before Swiss acedemics and military on 12th June 1975 successfully translated 30,000 of unknown text.

So, for example, when his system saw the sentence "Put the pyramid on the block in the box", it would immediately resolve the surface syntactic ambiguity of that command according to whether there was a block either under a pyramid or already in the box in the blocks scene that it understood.

More typical of an implicit response to MT problems were the systems of Schank, 1975a; Charniak, 1973; and Wilks, 1975, which, in their different ways, were concerned with the semantic representations, real world knowledge and inference rules needed to understand various aspects of every-day story-like sentences, and to produce deep representations for them, from which translations in another language could in principle be produced.

In the last few years the paradigm in AI and language understanding has itself shifted, largely in response to an argument of Minsky's (1975) that more complex knowledge structures were required than had been contained in any of the systems mentioned so far. He called these more complex structures frames, and argued that without the more specific knowledge of concepts that they expressed, language understanding would not be possible.

So far, as in (Schank, 1975b; Charniak, 1975) frames have been taken to be representations of stereotypical situations, such as the normal sequence of events in shopping in a supermarket. He argues that we can easily construct stories that will not be understood without such knowledge:

"John put some bacon in the basket in the supermarket, but then slipped a bar of chocolate off the shelf and into his pocket. When he got to the checkout his face went red and he said I didn't mean to take it." He might argue that one cannot refer the "it" correctly in that sentence, and so understand it or translate it into a suitably ~~gendordered~~ gendered language, without the sequential knowledge of what is and is not normal in a supermarket. There is still some unclarity about what precisely are the claims implied by the use of such frames (see Wilks, 1977b), but there is no doubt that they do represent a real form of language-related knowledge and can be seen as a new attempt to tackle the old MT problem of topic. So, for example, Schank has a program for using restaurant frames such that when it sees, say, the word "order" it will know that it is the word "order food", because it is encountered in a restaurant frames, and not the more general "order an object".

Schank has also supervised the construction of a program that reads stories into a frame format and then translates out the whole frame (including the stereotypical parts not actually mentioned in the original story) in a number of different languages.

In what remains of this paper I would like to sketch a proposal for the relevance of a rather different type of frame to MT: a static and not a dynamic frame to do with normal sequences of actions. It is directed towards another intractable problem of MT: of what to do when the input does not fit our semantic expectations. I will now turn to this, and then finally make some remarks about how far such AI proposals may take us in MT even when they eventually work as programs.

Meaning boundaries and knowledge structures

The remainder of the paper sketches how one might deal with extensions of word sense in a natural language understanding system (NLUS): that is to say, normal utterances that break preassigned selection, or preferences, restrictions. The proposals here extend the knowledge representation of the preference semantics NLUS (Wilks 1968, 1973, 1975) with pseudo-texts (PT) which are frame structures in the sense of (Minsky 1975), but which are also consistent with the general assumptions of this NLUS.

It is essential to see that extended use, in the sense of preference-violating use is the norm in ordinary language use, and so cannot be relegated in an NLUS to some special realm of treatment, as "performance" is in generative linguistics, nor neglected in a general consideration of MT. The following sentence is chosen, I promise you, at random from the front page of a daily newspaper: (The Times, 5-2-76):

(1) Mr. Wilson said that the line taken by the Shadow Cabinet, that a Assembly should be given no executive powers would lead to the break-up of the United Kingdom.

The sentence presents no understanding problems whatever to an informed reader, yet each of the four underlined entities violates the normal preference restrictions of an associated verb: lines, for example, would violate the normal physical object restriction on "take", and so on.

The process to be described in this paper is called projection: we shall show how sense descriptions for words can be rewritten, in preference-violating texts (as in 1), with the aid of the specific knowledge in PTs: part of the PT will be projected into the sense description for a word. So, for example, in (1) some detailed political knowledge in a PT for "United Kingdom" could show that a breaking of that entity could be caused, and we would then replace the sense description of "lead to" by one equivalent to "cause", thus overcoming the preference violation in "lead to the break-up" and providing a more appropriate sense description of "lead to" for analysis of the rest of this text.

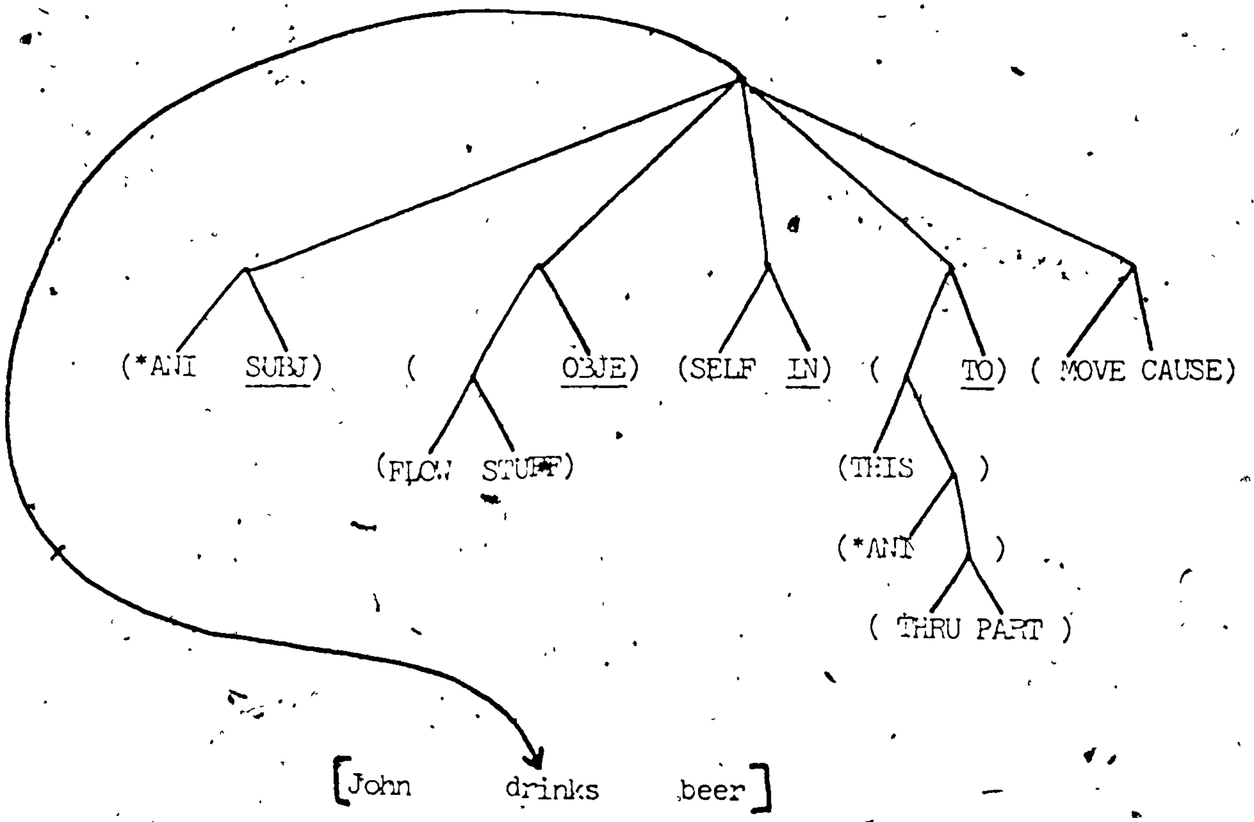
Brief recap of preference semantics

In previous papers I have described an NLUS in which rules operate on semantic word-sense descriptions to build up text descriptions. The rules that insert sense descriptions into text descriptions are what I have called "preferential": they seek preferred entities, but will accept the less preferred if necessary. A sense description for the action "drink" might be the formula:

fig. 1

This is a formal structure of semantic primitives expressing the meaning of the action (see King and Wilks 1977): that drinking is a CAUSing to MOVE, preferably done by an ANImate SUBJect (=agent) and to a liquid (FLOW STUFF), TO a particular

A SEMANTIC FORMULA FOR THE ACTION OF DRINKING.



THE ACTION FORMULA FOR DRINKING INSTALLED AT THE (CENTRAL) ACTION NODE OF A SEMANTIC TEMPLATE OF FORMULAS FOR "John drinks beer".

figure 1

ANImate aperture (THRU PART), and INto the SELF (=the animate agent). For short we will write (2) as [drink]. The text structures in the system are semantic templates (together with semantic ties between them): a template is a network of formulas, containing at least an agent, action and object formula. Thus the template for "The adder drinks water" will be written the+adder drinks water for short where the whole of (2) is in fact at the central (action) node.

The process of setting up the templates allows the formulas to compete to fill nodes in templates. Thus the formula for the (snake-)adder goes to the agent node in the template above in preference to the (machine-)adder because (2) specifies, by (*ANI SUBJ), that it prefers to be accompanied in a template by an animate agent formula. However, in the sentence:

(3) My car drinks gasoline

the available formula for the first template node, namely [car], is not for an animate entity, yet it is accepted because there is no competitor for the position. THE PURPOSE OF THIS PAPER IS TO SKETCH HOW THE SYSTEM MIGHT NOT MERELY ACCEPT SUCH A PREFERENCE-VIOLATING STRUCTURE FOR (3) BUT MIGHT ALSO INTERPRET IT.

An important later process is called extraction: template-like structures are inferred and added to the text representation even though they match nothing in the surface text. They are "deeper" inferences from the case structures of formulas in some actual template - where the case primitives are those underlined in (2). Thus, to the template for (3), we would add an extraction (in double

square parentheses in abbreviated form):

(4) [[gasoline in car]]

which is an inference extracted from the containment subformula of (2), (SELF IN). Analogous extractions could be made for each case primitive in each formula in the template for (3).

Since the programmed version of the system, reported in (Wilks 1975), a structural change (Wilks 1976a) has allowed a wider, and more specific, form of expression in formulas by allowing thesaurus items, as well as primitives, to function in them. No problems are introduced by doing this, provided that the thesaurus items are also themselves words in the dictionary, and so have their formulas defined elsewhere in their turn. One advantage of this extension is to impose a thesaurus structure on the whole vocabulary, and so render its semantic expression more consistent.

A thesaurus, like Roget, is simply an organisation of a vocabulary into semi-synonymous rows, which are themselves classified hierarchically under heads, and even more generally, sections. Thus under some very general section name MOVE (=motion) we would find heads, two of which might be ≠ engine and ≠ vehicle. The former might be the name of a row of actual types of engine:

(5) ≠ 525 engine: turbine, internal combustion, steam, ...

where the number simply indicates the sequence position of ≠ engine in the thesaurus. It is no accident that the most general section names like MOVE can be identified with the semantic primitives of the present system.

The organisation is imposed by requiring inclusion relations, between the formulas for word senses, corresponding to the thesaurus relations of the words. Thus, all the words in the row (5) would have a common subpart to their formulas, and that common subpart would be the dictionary formula for "engine", probably expressing in primitives no more than "a thing used by humans to perform some task, and self-moving in some way". If now thesaurus items can be inserted in formulas we may expect a formula for "car" at least as specific as:

(6)

fig. 2

Language boundaries and projection

Let us return to examples like (3) for which the system constructs a template even though it contains a violated preference, and ask what should an intelligent system infer in such a situation?* I would suggest that cars can be said to drink in virtue of something a system might already know about them, namely that they have a fluid (gas/petrol) injected into them, and they use that in order to run. That is to say, the program should have access to a sufficiently rich knowledge structure to express the fact that cars stand in a relation to a particular fluid, a relation that is of the "same semantic structure" as the relation in which a drinker normally stands to the thing drunk. All that may sound obvious, but how

---

* The sytem already deals with certain preference violations, such as those constituting the ergative case paradigm ("The hammer broke the sindow" see Wilks 1976b) and certain examples like "Joh got a shock", a class central to Riesbeck's thesis (see Schank (ed.) 1975).

A SEMANTIC FORMULA FOR "car".

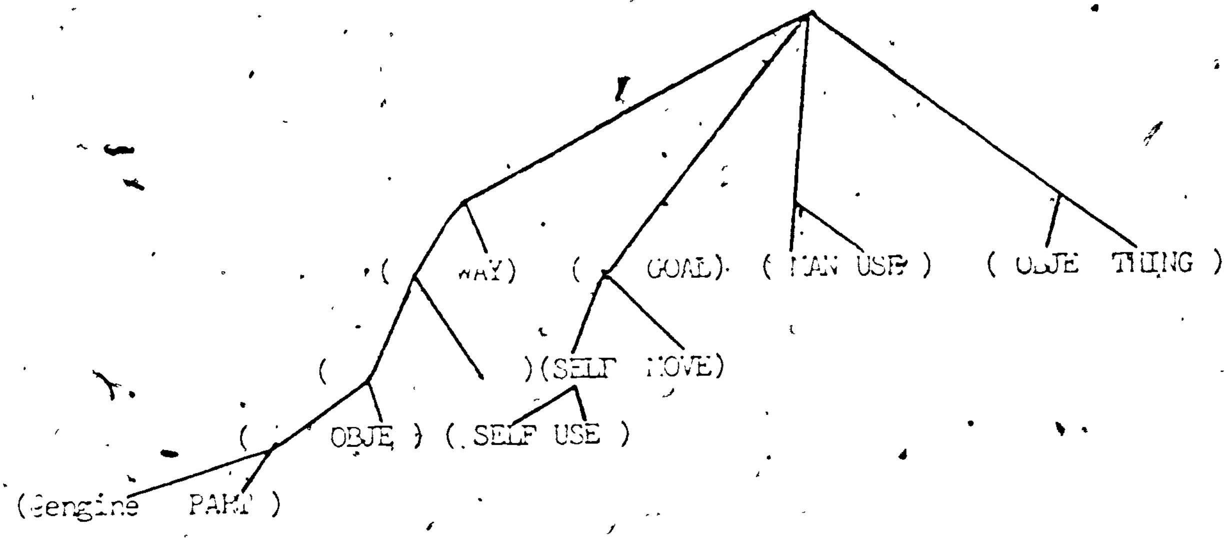


figure 2

else are we to account for the naturalness of (3), but the relative unnaturalness (and uninterpretability) of "My car chews gasoline", and, the more distant, "My car carves the Sunday roast". One upshot of these proposals is to distinguish plausible (with respect to a knowledge base) preference violation from the implausible.**

The procedural upshot of the above would be to replace at least one formula in the template for (3) with another, either constructed by rule*** or drawn from the knowledge structure itself, to be called a pseudo-text (PT). Let us now postulate that "car" points not only to (6), i.e. [car] but that [car] in turn points to:

(7)

fig. 3

This structure is called a pseudo-text because it is of just the same format as the text representations produced by the present NLUS. It can be extended to taste to express as much specific information about cars as is thought appropriate - given the parser for the present NLUS, it could even be input as a real text about

---

** An important aspect of the interpretation of (3) is idiomatic, namely that the car uses a lot of gas/petrol. This aspect of the meaning is beyond this, or I suspect any, general inference procedure.

*** In a fuller version of this paper (Wilks, 1977) I describe the relation of this work to attempts, such as (Givon 1967), to give general rules for projection: rules operating on the dictionary and independent of contexts of use.

PART OF A PSEUDO-TEXT FOR "car".

	[(MAN)	@inject	@liquid		] a human injects a liquid
INST	[(MAN)	(USE)	@tube		] using a tube
TO	[@liquid	(IN)	*		] liquid is in the car
	[IC engine	(USE)	@liquid		] the ICengine uses the liquid
CAUSE	[*	(MOVE)	□		] the car moves
GOAL	[((* IN)MAN)	(MOVE)	□	CAUSE	] the human in the car moves
	[((* IN)MAN)	@turn	@wheel		] the human turns the wheel
INST	[(MAN)	(USE)	(MAN PART)		] the human uses a selfpart
TO	[*	(TO)	(NOTSAME POINT)		] the car moves to a new spatial point.

and so on..........

figure 3

cars. The representation consists of the templates (explained loosely at the right), together with the (self-explanatory) case and cause ties between them. In the templates, □ dummy and * denotes the formula [car] that points to this object (7). The ≠ prefixed items are thesaurus items, though "IC engine" is simply a specific dictionary word pointing to its own formula - specificity is thus a matter of taste. So, for example, the thesaurus head ≠ liquid could be replaced by the more explicit "gasoline". Items in round parentheses remain in primitive form. It will be clear that the same information can be expressed in a number of different ways, and at different levels of generality; though the spirit of (Minsky 1975) suggests that they should be as specific as possible. The intention here is that THE PROCESSES THAT OPERATE ON SUCH ENTITIES AS (7) SHALL BE IDENTICAL WITH THOSE THAT MANIPULATE REPRESENTATIONS DERIVED FROM INPUT TEXTS. The approach is thus the reverse of the conventional one: we seek to assimilate knowledge structures to text structures, rather than the reverse, on the grounds that the representation of language is the difficult task, and that the representation of knowledge as such makes no sense apart from that.

We should note, too, that just as the thesaurus structure imposes a containment relation on the formulas of co-row-member words, so it also imposes a hierarchical relationship on PTs: that for vehicle, for example, will be a less specific version of (7). Further up the thesaurus would be PTs for high-level sections: that for MAN would be highly complex, for example. But note there is no "inheritance of property" problem in this system: the formula for "amputee" would have head MAN and would specify the loss of limbs. Any inherited pseudo-text from MAN - asserting "two legs" - would be modified by [amputee] .

The system now uses (7) to make a projection, so as to derive an interpretation for (4), by seeking, in (7) templates matching the source template [my+car drinks gasoline] : namely the first and fourth lines of (7). The first match is in virtue of the similarity of [drink] and [≠inject] - based on the expression in primitives, as in (2), of causing a liquid to be in an entity of the same type as the agent. This would allow us to confirm by projection, the "humanness of the drinker", that has already been noted by earlier extraction* routines, extracting out from the drink formula (2) independently of the PT (7). However, no projection is made at this stage onto [car], (though it might be later in the face of a sentence after (4) like "His thirst is never slaked", that confirms the humanness projection) because in the case of violations of the preferences of actions, like "drink" in (4), the system ALWAYS PREFERS TO MAKE A PROJECTION ONTO THE ACTION ITSELF IF IT CAN. The strong match detected is between the above template for (3) and the fourth line of (7) in virtue of the containment of [≠engine] in [car], and by [≠liquid] of [gasoline], which is evident in the formulas themselves. This results in the projection of the action node of the fourth line of (7), namely [use], onto [drink] in the template for (3). This projection is taken to be strongly confirmed by the match with the first line of (7), and is considered to carry over more sense than any alternative projection. The confirmation (of the match of the fourth line of (7) by the first line) is necessary here, because [my+car leaks gasoline] would also match the fourth line, but no such projection would be appropriate. Conversely, no projection could be made for "My car drinks mud" from the fourth line, even with the confirmation of the first. The general

---

* extractions, it will be seen, differ from projections in that they produce new template-like entities, rather than, as here, replacing formulas inside existing templates.

rule for action projections then is: SEEK A PSEUDO-TEXT, FOR AGENT OR OBJECT, WITH A TEMPLATE MATCHING ON AGENT AND OBJECT NODES. PROJECT THIS GENERALLY IF THERE IS ALSO A PSEUDO-TEXT TEMPLATE MATCH TO THE ACTION ITSELF, FOR ANOTHER TEMPLATE IN THE SAME PSEUDO-TEXT.

We may note in passing three interesting developments of the above suggestion. First consider the more complex example presented by a recent headline: (8) Britain tries to escape Common Market.

Clearly, some projection would be appropriate here, of humanness onto the country, and perhaps even "prisonlikeness" onto the formula for the Common Market. These might be drawn from the formula for "escape" alone, by extraction and without recourse to the pseudo-texts for either of the entities. Even if we did consult those entities, we would find a historical account of Britain joining, but not of leaving. In such circumstances mere facts are not enough, even when highly structured. We might conceivably be able to project some notion [disassociate] onto [escape], from the "Britain pseudo-text", given some new matching criterion that placed relevance above negation in such cases (i.e. would match [escape] with [associate] or [join].)

Secondly, we might consider the problems presented by an example like:

(9) I see what you mean.

Here the last clause breaks the preference expressed in [see] for a physical object. A system procedure will present the actual object of (9) to the top-level

template simply as the primitive SIGN (the primitive for symbols and intensional representations of them) which has been obtained, by extraction, from the preferred object in [mean]. Thus the system is effectively dealing with the template sequence [I see (SIGN)][you mean (SIGN)]. But what could we expect as a pseudo-text for something as general as SIGN, so as to use the above procedures to project onto [see]. If we take advantage of the hierarchical nature of the thesaurus, we might expect pseudo-texts at the very top level, associated with the section names - pure primitives like SIGN - just as specific pseudo-texts are associated with the lowest level items in the thesaurus - row members like "car". The pseudo-text for a primitive like SIGN would be wholly "core structural": it would consist of no more than primitive concatenations, in template form, like MAN THINK SIGN*, the most general thing that can be said about what is normally done to signs. However, even something as general as this might suffice to project THINK correctly onto [see]. The interesting generality would come from using <u>exactly the same projection procedures</u> on the most general pseudo-texts like this, as on the most specific, like (7).

Thirdly, and this treated at length in Wilks (1977), we can consider a quite different type of projection for phrases like:

(10) a toy lion

This comes from a much discussed class of examples ("plastic flower", "stone horse" etc.), where an obvious projection mechanism is to replace the head of the formula for the noun (BEAST in [lion] in (10)) by the preferred object of predication in the qualifier - here *PHYSOB in [toy]. This would be a very

---

* those familiar with the system of Wilks (1968, 1965a etc.) will remember that these are the "bare template" structures actually used to obtain the initial template match. The suggestion here is that the "knowledge-aspect" of these highly-general structures is to be found as the pseudo-texts of primitives - as the latter function right at the top of the conceptual hierarchy imposed by the thesaurus.

limited and general class of projections, not requiring access to PTs, but which might still provide a "projected formula" appropriate for examples like:

(11) The cat walked round the toy lion.

Then he came back and sniffed it.

where we might be helped to refer "he" and "it" correctly by the new, projected, formula [lion] whose head was no longer BEAST, and which could therefore no longer be the reference of "he" as a real lion would be.

A more radical and interesting development would be the construction of "PT repacking functions" specific to certain qualifiers. Thus, for example, such a function for "toy", if faced with the phrase "toy car" might repack (7) using a general rule to delete all constituent templates based on the action USE, as well as all those that are at end of a GOAL tie, since toy cars cannot, normally, serve human needs, uses and purposes.

Postscript

The above suggestions are, as should be clear, only in the pre-program stage, but they will be implemented.

What I have tried to suggest in this paper is that AI language programs do bear upon the traditional difficulties of MT, and often do so more directly than conventional linguistic theories, with their preoccupation with well-formedness, and with delimiting the class of all utterances of a language.

I have given the impression perhaps that all AI programs are concerned with what could be called stratospheric considerations: the solution of the most general problems of language and understanding. That would be unfair: there is a number of more task-oriented projects under construction, attempting to limit vocabulary and world knowledge to very limited domains, such as plumbing repair, say, so as to produce concrete results while at the same time appealing to very general philosophical principles (see Levin & Moore, 1976).

What all the AI projects, of whatever level, have in common is an appeal to very general knowledge and principles, coupled to the claim that MT work must take account of these if it is ever to achieve any generality and reliability. The reply to this claim, from experience with projects like SYSTRAN, is that the examples that make AI these points are artificial and/or rare, and they can be ignored for practical purposes. This is clearly an empirical dispute and open to test, which is what makes the present situation interesting as I remarked at the beginning.

That much does depend on one's choice of examples can be seen by returning to those of the beginning: Bar-Hillel's "slow neutrons and protons" should be amenable to treatment by an expert "atomic physics frame", one no more open to the charge of "ad hocness" than is human's knowledge of physics itself. But with the old favourite "Time flies like an arrow", things are not so clear. In terms of what I called preferences, it may well be that the desired reading (where time does the flying) satisfies no more semantic preferences than, say, the reading where the flies like certain objects. Moreover, it is hard to imagine that any topic-determining frame could help here---one would hardly expect this slogan in any frame about time, except as an arbitrary addition. Nothing that has come from recent "speech act" theorists in linguistics and philosophy seems likely to help either. Perhaps, the only explanation of our competence with this sentence is that we read it off a list of cliches for which we have the assigned readings: a sad conclusion for all theoretically-motivated work, and an awful fate for a long cherished example.

References

ALPAC: Automatic Language Processing Advisory Committee (1966) Language and Machines. Publication No. 1416. Washington, D.C. National Academy of Sciences.

Bar-Hillel, Y. (1962) The future of machine translation. Times Literary Supplement, London : Times Newspapers. April 20th 1962.

Bobrow, D. & Winograd, T. (1977) KRL - an overview of a knowledge representation language. Cognitive Science, 1, 3-46.

Bresnan, J. (1976) Towards a realistic model of transformational grammar. Unpublished mss.

Bruderer, H. (in press) Handbook of Machine Translation and Machine-Aided Translation. Amsterdam : North Holland.

Chandioux, J. (1976) METEO, In Hays - Mathias (eds). pp 27-37.

Charniak, E. (1973) Jack and Jane in search of a theory of knowledge. In Proc. Third Internat. Joint Conf. on A.I. Menlo Park, Calif. : Stanford Research Inst. 115-124.

Charniak, E. (1975) Organization and Inference. In Proc. Theoretical Issues in Natural Language Processing. Cambridge, Mass: M.I.T. 105-114.

Etat des activités multilingues en matière d'information scientifique et technique : Vol 1, Rapport Final. (1976 Bureau M. Brussels : van Dijk.

Hays, D. and Mathias, J. (eds) (1976) Proc. FBIS Seminar on Machine Translation. Amer. Jnl. Computl. Linguistics 46 1976

Fillmore, C. (1971) On a fully developed system of linguistic description. In Lehmann and Stachowitz (eds) pp. 77-94.

Givon, T. (1967) The structure of ellipsis. mimeo. Santa Monica. Calif.: Systems Development Corp.

Grice, H. (1967) Logic and Conversation. Unpublished mss.

King, M. & Wilks, Y. (1977) Semantics, Preference and Inference. Geneva: Institute for Semantic and Cognitive Studies.

Levin, J. and Moore, J. (1976) Dialogue Games. In Proc. A.I.S.B. Conference. Edinburgh : Dept. of Artificial Intelligence.

Lehmann, W. and Stachowitz, R. (eds) (1971) Feasibility study on fully automatic high-quality translation, Report RADC-TR-71-295. Rome, N.Y. : Rome A.F. Development Center.

Minsky, M. (ed) (1968) Semantic Information Processing. Cambridge, Mass. : M.I.T. Press.

Minsky, M. (1975) A framework for representing knowledge. In Winston (ed) The Psychology of Computer Vision. New York : McGraw Hill 211-277.

Schank, R. (ed) (1975a) Conceptual Information Processing. Amsterdam : North Holland.

Schank, R. (1975b) Using knowledge to understand. In Proc. Theoretical Issues in Natural Language Processing. Cambridge, Mass. : M.I.T. 67-77.

Steiner, G. (1975) After Babel : Aspects of Language and Translation. London : Oxford U.P.

Wilks, Y. (1968) Computable Semantic Derivations. (As given).

Wilks, Y. (1973) An artificial intelligence approach to machine translation. In Schank and Colby (eds) Computer Models of Thought and Language. San Francisco : Freeman. 101-129.

Wilks, Y. (1975) A preferential, pattern-matching, semantics for natural language understanding, Artificial Intelligence. 6, 53-76.

Wilks, Y. (1976a) De Minimis : the archaeology of frames. In Proc. AISB Conference Edinburgh, Dept. of A.I. : 133-142.

Wilks, Y. (1976b) Processing Case. Amer. Jnl. Comput. Ling. 56.

Wilks, Y. (1977a) Making Preferences more active. mimeo Edinburgh : Dept. of A.I., memo no. 32.

Wilks, Y. (1977b) Frames, scripts, stories and fantasies. Pragmatics Microfiche, 3.

Winograd, T. (1972) Understanding Natural Language. Edinburgh : Edinburgh U.P.